# Functions of Bounded Variation on the Classical Wiener Space and an Extended Ocone-Karatzas Formula

M. Pratelli[a], D. Trevisan[b,*]

*[a]Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo, 5, 56127, Pisa, Italy*
*[b]Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56126, Pisa, Italy.*
*Mobile: +39 331 2899761*

## Abstract

We prove an extension of the Ocone-Karatzas integral representation, valid for all $BV$ functions on the classical Wiener space. We establish also an elementary chain rule formula and combine the two results to compute explicit integral representations for some classes of $BV$ composite random variables.

*Keywords:* $BV$ functions, Wiener space, Ocone-Karatzas formula

## 1. Introduction

Functions of bounded variation ($BV$) in a Gaussian Banach space setting were first investigated by M. Fukushima and M. Hino in [1] and [2], using techniques from the theory of Dirichlet forms. More recently, L. Ambrosio and his co-workers gave an alternative approach, in [3], by adapting techniques from geometric measure theory.

The most important example of an infinite-dimensional Gaussian space is given by the classical Wiener space $(\Omega, \mathcal{A}, \mathbb{P})$, i.e. the space of trajectories of the Wiener process. This was the setting where the Malliavin calculus was originally developed and still most of its applications are formulated. Since $BV$ functions generalize Malliavin differentiable functions, we specialize here the general results valid for all Gaussian spaces, work on explicit examples and consider new problems which appear naturally, in connection with stochastic analysis. Thus, all the results given here are formulated in this setting, though some of them are certainly valid in any abstract Wiener space.

One of the aims of this paper is to study how the theory of $BV$ functions can be a useful additional tool, even when dealing with classical problems. We believe that the extension of the Ocone-Karatzas formula, Theorem 16, can be

---

*Corresponding author
Email addresses:* `pratelli@dm.unipi.it` (M. Pratelli), `dario.trevisan@sns.it` (D. Trevisan)

regarded as the best example, in this direction, among those presented in this paper.

Roughly speaking, a real function $f$ defined on the classical Wiener space is $BV$ if it admits an $L^2(0,T)$-valued measure $Df$ which plays the role of a Malliavin derivative, so that an integration-by-parts identity holds true. Generalizing the situation of differentiable functions, where the Malliavin derivative is a stochastic process, $Df$ can be seen as a measure on the product space $(\Omega \times [0,T], \mathcal{A} \otimes \mathcal{B}(0,T))$ and processes can be integrated with respect to it.

Notably, $Df$ can be not absolutely continuous with respect to $\mathbb{P} \otimes \lambda$ (where $\lambda$ is the Lebesgue measure). However, if we introduce the strictly predictable $\sigma$-algebra $\mathcal{P}$, which is slightly smaller than the usual $\sigma$-algebra of predictable sets in $\Omega \times [0,T]$, then $Df$ restricted to $\mathcal{P}$ becomes absolutely continuous with respect to $\mathbb{P} \otimes \lambda$. Moreover, if $H = (H_s)_{0 \leq s \leq T}$ is a version of the density, then

$$f = \mathbb{E}[f] + \int_0^T H_s dW_s.$$

This is, informally, the content of Theorem 16, that is our extension of the classical Ocone-Karatzas formula, originally proved in [4], which identifies the integrand in the Itô representation of a random variable in terms of its Malliavin derivative. In the differentiable case, one usually writes $H_s = \mathbb{E}(\partial_s f | \mathcal{F}_s)$, for $\lambda$ a.e. $s \in [0,T]$. Proposition 13 shows that an similar result holds true in the $BV$ case, which can be useful when dealing with concrete cases, although from a higher point of view the process $H$ should be considered as the (dual) predictable projection of the measure $Df$.

$BV$ functions can be helpful when dealing with a composite function $f = \phi \circ g$, where $g$ is a real random variable, differentiable in the Malliavin sense, and $\phi$ is a Euclidean $BV$ function. In such a case, as in the classical theory of Euclidean $BV$ functions, many problems for general functions can be reduced to the case of indicator functions of level sets $\{x > g\}$, which can be shown to be $BV$, under certain assumptions. Moreover, an explicit chain rule formula can be obtained (Theorems 8 and 10). We remark here that this chain rule is still very distant from the deep results which can be obtained in the Euclidean setting, but it can be useful when dealing with applications.

Indeed, as an application, we combine the extended Ocone-Karatzas formula and the chain rule, to obtain explicit representations for some functionals of the Wiener process (Propositions 20 and 22). While the former result is well-known, nevertheless we believe that the theory of $BV$ functions provides, we believe, a clear proof without advanced technical results, such as the theory of distributions on Wiener spaces.

This paper is organized as follows: in Section 2, we fix some notation and provide the definition of $BV$ functions together with the main approximation result, Theorem 6, without proof. Other technical results, e.g. on the Orlicz space $L \log^{1/2} L$, are collected, for the convenience of the reader. In Section 3, we investigate a chain rule for a special class of $BV$ functions. Here, we use an approximation result for Euclidean $BV$ functions, though an elementary proof

2

is given for the special case of indicator functions of level sets. In Section 4, the extended Ocone-Karatzas formula is established, after some remarks on the notion of predictability. In Section 5, applications and examples are discussed.

## 2. Notation and preliminary results

### 2.1. Malliavin calculus

Let us fix $T > 0$ and consider the classical Wiener space $(\Omega, \mathcal{A}, \mathbb{P})$, where $\Omega = C_0([0, T])$ is the space of real continuous functions $\omega$, with $\omega(0) = 0$, $\mathcal{A}$ is the $\sigma$-algebra of Borel sets of $\Omega$, and $\mathbb{P}$ is the Wiener measure on $\mathcal{A}$. By definition, with respect to the probability measure $\mathbb{P}$, the canonical process $(W_t)_{0 \leq t \leq T}$, given by $W_t(\omega) = \omega(t)$, is a Wiener process starting from the origin.

For a complete exposition of the Malliavin calculus on the classical Wiener space, we refer to [5]. Here, we recall some basic facts, together with some minor changes in the notation, which turn out to be more convenient, when dealing with $BV$ functions.

We write $L^2(0, T)$ for $L^2([0, T], \lambda)$, where $\lambda$ is the Borel-Lebesgue measure, restricted to the interval $[0, T]$. We write $\langle k_1, k_2 \rangle_2$ for the scalar product between $k_1, k_2 \in L^2(0, T)$.

The classical Cameron-Martin space $H_0^1 \subset \Omega$ is the space of real continuous functions $h$ on $[0, T]$, such that, for some $k \in L^2(0, T)$,

$$h(t) = \int_0^t k(s)\, ds,$$

for every $0 \leq t \leq T$. Therefore $k = h'$, $\lambda$-almost everywhere. Endowed with the scalar product $\langle h_1, h_2 \rangle_{H_0^1} = \langle h_1', h_2' \rangle_2$, $H_0^1$ is a Hilbert space isomorphic to $L^2(0, T)$.

The Wiener integral construction allows us to identify $L^2(0, T)$ with a subspace of $L^2(\Omega, \mathcal{A}, \mathbb{P})$, and therefore $H_0^1$ with the same subspace, with the correspondence $h \mapsto W(h') = \int_0^T h'(s)\, dW_s$.

For $n \geq 1$, given an $n + 1$-uple of times $J = (t_0, \ldots, t_n)$, with $0 \leq t_0 < \ldots < t_n \leq T$, we define

$$\Delta_J W = \left( W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}} \right). \tag{1}$$

We will often omit the subscript $J$ and simply write $\Delta W$.

A smooth function $f$ is a real function of the form $f = \phi(\Delta_J W)$ for some $J$ as above and some $\phi \in C_b^1(\mathbb{R}^n)$. We remark that $\mathbb{P}$ plays no role in this definition, thus allowing us to consider different measures on $(\Omega, \mathcal{A})$.

It can be shown that any $f = \phi(\Delta W)$, with a bounded continuous $\phi$, is the pointwise limit of a uniformly bounded sequence of smooth functions: indeed, it is sufficient to approximate $\phi$ with a sequence of smooth functions on $\mathbb{R}^n$. Then, by the monotone class theorem, or another equivalent approximation argument, we obtain that, given a finite positive measure $\mu$, smooth functions are dense in

3

$L^p(\Omega, \mathcal{A}, \mu)$, for every $1 \leq p < \infty$. In particular, this holds true for $\mu = \mathbb{P}$. We will often write $L^p(\mathbb{P}) = L^p(\Omega, \mathcal{A}, \mathbb{P})$.

By definition, the Malliavin derivative of a smooth function $f = \phi(\Delta W)$ is the application $\nabla f : \Omega \to L^2(0, T)$,

$$\nabla f(\omega) = \sum_{i=1}^{n} \partial_i \phi(\Delta W(\omega)) I_{]t_{i-1}, t_i]} \in L^2(0, T).$$

The Malliavin derivative is well-defined, since

$$\langle \nabla f, h' \rangle_2 = \sum_{i=1}^{n} \partial_i \phi(\Delta W) [h(t_i) - h(t_{i-1})] = \partial_h f.$$

Given $h \in H_0^1$ and smooth functions $f, g$, if we write $\partial_h^* g = \partial_h g - gW(h')$, the integration-by-parts formula holds,

$$\mathbb{E}[(\partial_h f) g] = -\mathbb{E}[f \partial_h^* g].$$

It follows that, for every $p \geq 1$, the linear operator $\nabla$ is well-defined on a dense subset of $L^p(\mathbb{P})$, with values in $L^p(\mathbb{P}; L^2(0, T))$, and closable. We denote the domain of its closure by $\mathbb{D}^{1,p}$.

By an explicit approximation, $\nabla W(k) = k$, for every $k \in L^2(0, T)$. Therefore, our definition of smooth functions provides a construction of the Malliavin derivative, equivalent to that in [5].

The fact that the Malliavin derivative of a random variable can be identified with a process is a consequence of the following elementary result, which can be easily proved by a density argument. We prefer to state it as a lemma, since it will be used again when dealing with $BV$ functions.

**Lemma 1.** *Given a positive measure $\nu$ on $(\Omega, \mathcal{A})$, there is a linear continuous immersion*

$$L^1(\Omega, \nu; L^2(0, T)) \to L^1(\Omega \times [0, T], \mathcal{A} \otimes \mathcal{B}([0, T]), \nu \otimes \lambda)$$

*that maps every $F$ to a process $F^\nu$ such that, $\nu$ almost everywhere, the function $t \mapsto F_t^\nu(\omega)$ coincides with $F(\omega)$, $\lambda$ almost everywhere. Moreover,*

$$\int_{\Omega \times [0, T]} |F^\nu| \, d(\mu \otimes \lambda) \leq T^{1/2} \int_\Omega |F|_2 \, d\nu.$$

Therefore, given $f \in \mathbb{D}^{1,1}$, we write $(\partial_t f)_{0 \leq t \leq T}$ for the process $\nabla f^\mathbb{P}$ given by the lemma above. For every $k \in L^2(0, T)$,

$$\langle \nabla f, k \rangle_2 = \int_0^T \partial_s f \, k(s) \, ds, \quad \mathbb{P} \text{ a.s.}$$

4

### 2.2. The space $L \log^{1/2} L$

We write $A_{1/2}$ for the real continuous convex function

$$x \mapsto A_{1/2}(x) = \int_0^{|x|} \log^{1/2}(1+s)\, ds.$$

By definition, $L \log^{1/2} L = L \log^{1/2} L(\mathbb{P})$ is the vector space of real random variables $X$, such that, for some $k > 0$, $A_{1/2}(X/k) \in L^1(\mathbb{P})$. Endowed with the Luxembourg norm

$$\|X\|_{L \log^{1/2} L} = \inf \left\{ k > 0 : \mathbb{E}\left[ A_{1/2}(X/k) \right] \le 1 \right\},$$

it is a particular case of an Orlicz space, and therefore a Banach space.

Orlicz spaces generalize $L^p$ spaces and, if the function which defines the norm does not grow too fast, many properties can be stated and proved in exactly the same way. $L \log^{1/2} L$ is such an example of slow growth, since for every real $0 < k_1 < k_2$ and every real $x$,

$$\log^{1/2}(1 + x/k_1)/k_1 < (k_2/k_1)^{3/2} \log^{1/2}(1 + x/k_2)/k_2.$$

It follows that

$$\mathbb{E}\left[ A_{1/2}(X/k_1) \right] \le (k_2/k_1)^{3/2} \mathbb{E}\left[ A_{1/2}(X/k_2) \right],$$

and so, if $X \in L \log^{1/2} L$ then, for every $k > 0$, $A_{1/2}(X/k)$ is integrable.

Given $X \in L \log^{1/2} L$ and a real random variable $Y$ with centred normal law then, for some $C(Y)$,

$$\mathbb{E}\left[ |XY| \right] \le C(Y) \|X\|_{L \log^{1/2} L}. \tag{2}$$

This follows, for example, from Young's inequality for $A_{1/2}$,

$$|xy| \le A_{1/2}(x) + \int_0^{|y|} \left( e^{t^2} - 1 \right) dt \le A_{1/2}(x) + |y| e^{y^2},$$

and taking expectation, with $x = X/\|X\|_{L \log^{1/2} L}$ and $y = Y/2\sqrt{\mathbb{E}[Y^2]}$.

In the context of Malliavin differentiable functions, an important consequence of the result above is that, given a smooth function $g$ and $h \in H_0^1$, $\mathbb{E}[f \partial_h^* g]$ is well-defined whenever $f \in L \log^{1/2} L$.

*Remark* 2. Another fundamental fact, of which we will make implicit use, since it is partially contained in Theorem 6, is the existence of a continuous embedding of $\mathbb{D}^{1,1}$ into $L \log^{1/2} L$ (see [2], Proposition 3.2). We remark that, in turn, this is a consequence of the Gaussian isoperimetric inequality (see [6]).

We turn now to a technical result concerning the convergence of closed martingales in the space $L \log^{1/2} L$, which will be used in the proof of Theorem 16.

**Lemma 3.** *Given a discrete filtration $\mathcal{G} = (\mathcal{G}_n)_{n \geq 1}$, with $\bigvee_n G_n = \mathcal{A}$, and a $\mathcal{G}$-martingale $(M_n)_{n \geq 1}$, closed by a random variable $M \in L \log^{1/2} L$ (so that for every $n \geq 1$, $M_n = \mathbb{E}[M|\mathcal{G}_n]$), then the sequence $(M_n)_{n \geq 1}$ converges to $M$ in $L \log^{1/2} L$.*

*Proof.* It is well known that $(M_n)_{n \geq 1}$ converges $\mathbb{P}$ a.s. to $M$ and therefore in probability. We fix $\epsilon > 0$ and we show that

$$\limsup_n \mathbb{E}\left[A_{1/2}\left(\frac{M_n - M}{\epsilon}\right)\right] < 1.$$

Since $A_{1/2}(0) = 0$, by continuity, there exists some $\delta > 0$ such that $A_{1/2}(x) \leq 1/2$ if $|x| \leq \delta$. We write $B(n) = \{|M_n - M| > \delta\epsilon\}$, so that

$$\mathbb{E}\left[A_{1/2}\left(\frac{M_n - M}{\epsilon}\right)\right] \leq \frac{1}{2} + \mathbb{E}\left[I_{B(n)}A_{1/2}\left(\frac{M_n - M}{\epsilon}\right)\right].$$

Moreover, $\lim_n \mathbb{P}(B(n)) = 0$. By the elementary properties of $A_{1/2}$,

$$A_{1/2}\left(\frac{M_n - M}{\epsilon}\right) \leq \frac{1}{2}A_{1/2}\left(\frac{2M}{\epsilon}\right) + \frac{1}{2}A_{1/2}\left(\frac{2M_n}{\epsilon}\right).$$

$M$ belongs to $L \log^{1/2} L$, so the first summand above is integrable and

$$\limsup_n \mathbb{E}\left[I_{B(n)}A_{1/2}\left(\frac{2M}{\epsilon}\right)\right] = 0.$$

Since $A_{1/2}$ is convex, the second summand above is $\mathbb{P}$ a.s. not greater than $1/2\mathbb{E}\left[A_{1/2}(2M/\epsilon)|\mathcal{G}_n\right]$, by Jensen's inequality for conditional expectations. Again, $A_{1/2}(2M/\epsilon)$ is integrable so $\left(\mathbb{E}\left[A_{1/2}(2M/\epsilon)|\mathcal{G}_n\right]\right)_{n \geq 1}$ is a $\mathcal{G}$-martingale closed in $L^1$ and therefore uniformly integrable. We conclude that

$$\limsup_n \mathbb{E}\left[I_{B(n)}A_{1/2}\left(\frac{2M_n}{\epsilon}\right)\right] \leq \limsup_n \mathbb{E}\left[I_{B(n)}\mathbb{E}\left[A_{1/2}\left(\frac{2M}{\epsilon}\right)|\mathcal{G}_n\right]\right] = 0.$$

$\square$

### 2.3. $L^2(0,T)$-valued measures

We write $\mathcal{M} = \mathcal{M}\left(\Omega; L^2(0,T)\right)$ for the space of $L^2(0,T)$-valued $\sigma$-additive measures on $(\Omega, \mathcal{A})$, with finite total variation $|\mu|$. We recall that $|\mu|$ is a measure on $(\Omega, \mathcal{A})$, given by

$$|\mu|(A) := \sup\left\{\sum_{n \geq 1}|\mu(A_n)|_2 : A = \bigcup_{n \geq 1}A_n\right\} < \infty,$$

where the supremum runs along every countable measurable partition of $A$.

By the polar decomposition theorem, given $\mu \in \mathcal{M}$, there exists an $L^2(0,T)$-valued measurable application $\sigma$, with $|\sigma|(\omega) \leq 1$, for all $\omega$, such that, for every measurable set $A$, and every $k \in L^2(0,T)$,

$$\langle \mu(A), k \rangle_2 = \int I_A \langle \sigma, k \rangle_2 \, d|\mu|.$$

The member on the right clearly defines a real measure, which we denote by $\langle \mu, k \rangle_2$. The decomposition above allows us to integrate $L^2(0,T)$-valued applications. We define

$$\int \langle F, d\mu \rangle_2 = \int \langle \sigma, F \rangle_2 \, d|\mu|,$$

if $\langle \sigma, F \rangle_2 \in L^1(|\mu|)$. Since such applications can be seen as processes, we identify $\mu$ with a real measure on the product space. More precisely, we apply Lemma 1 with $\nu = |\mu|$ and consider the measure $\tilde{\mu} = \sigma^{|\mu|}.(|\mu| \otimes \lambda)$. Given a process $F = f I_{]s,t]}$, with smooth $f$ and $0 \leq s \leq t \leq T$, then

$$\int \langle F, d\mu \rangle_2 = \int f d \langle \mu, I_{]s,t]} \rangle_2 = \int_\Omega f(\omega) \left[ \int_s^t \sigma^{|\mu|}(\omega, r) \, dr \right] |\mu|(d\omega).$$

*2.4. BV functions*

The following condition shows that the integration-by-parts formula plays a central role in the theory of Malliavin calculus: $f \in \mathbb{D}^{1,p}$, with $1 \leq p < \infty$, if and only if $f \in L^p(\mathbb{P})$ and there exists some $F \in L^p\left(\Omega, \mathbb{P}; L^2(0,T)\right)$ such that, for every smooth $g$ and every $h \in H_0^1$,

$$\mathbb{E}\left[ \langle F, h' \rangle_2 \, g \right] = -\mathbb{E}\left[ f \partial_h^* g \right]. \tag{3}$$

In such a case, $F = \nabla f$.

We can read the left member in (3) as the integral of $g$, with respect to the real measure $\langle F, h' \rangle.\mathbb{P}$. Informally, a function $f$ is said to be of bounded variation, if we require only that there exists a measure, such that the same condition is satisfied.

**Definition 4.** A real function $f$ is said to be of *bounded variation* $(BV)$, with respect to $\mathbb{P}$, if $f \in L \log^{1/2} L(\mathbb{P})$, and there exists a measure $Df \in \mathcal{M}$, such that, for every $h \in H_0^1$ and every smooth $g$,

$$\int g \, d \langle Df, h' \rangle_2 = -\mathbb{E}\left[ f \, \partial_h^* g \right].$$

The quantity $|Df|(\Omega)$ is called the *total variation* of $f$. The real measure $\langle Df, h' \rangle_2$ is denoted with $D_h f$.

*Remark* 5. In the article [3], it is required for a smooth function to be of the form $\phi(W(k_1), \ldots, W(k_n))$, where $\phi$ is smooth and each $k_i \in L^2(0,T)$ is a function of bounded variation on $[0,T]$, so that $W(k_i)$ is linear and continuous on $\Omega$. Therefore, although the definitions are formally identical, the class of $BV$ functions introduced above could be larger than that considered there. However, by an approximation argument, it can be shown that they coincide.

Any function $f \in \mathbb{D}^{1,1}$ is $BV$, with $Df = \nabla f.\mathbb{P}$, because the integrability condition follows from the continuous embedding of $\mathbb{D}^{1,1}$ in $L \log^{1/2} L (\mathbb{P})$ (see Remark 2). Actually, a $BV$ function $f$ admits a Malliavin derivative if and only if $|Df|$ is absolutely continuous with respect to $\mathbb{P}$.

As a consequence of the general results about $L^2 (0, T)$-valued measures, in all what follows we will identify the measure $Df$ with a measure on the product space $\Omega \times [0, T]$. We can even define a $BV$-analogue of the process $(\partial_t f)_{0 \leq t \leq T}$: given a version $\sigma$ of the density of $Df$ with respect to $|Df| \otimes \lambda$, we define $D_t f = \sigma(t) . (|Df| \otimes \lambda)$. The family of measures $(D_t f)_{0 \leq t \leq T}$ is defined up to $\lambda$-neglegible sets.

We consider now the approximability of $BV$ functions with regular functions. To make a comparison with the differentiable case, we recall that, for $p > 1$, given a sequence $(f_n)_{n \geq 1}$ in $\mathbb{D}^{1,p}$, convergent to some $f$ in $L^p (\mathbb{P})$ and such that $(\nabla f_n)_{n \geq 1}$ is bounded in $L^p (\Omega, \mathbb{P}; L^2 (0, T))$, then $f \in \mathbb{D}^{1,p}$ (see [5], Lemma 1.2.3, p. 28, for the case $p = 2$). Such a conclusion does not hold, for $p = 1$: however, the next theorem shows that $f$ must be a $BV$ function, and all $BV$ functions can be obtained with a similar approximation. Proofs can be found in [2], Theorem 3.7, or in [3], Theorem 4.1.

**Theorem 6.** *Given $f \in L^1 (\mathbb{P})$, the following conditions are equivalent:*

1. *$f$ is of bounded variation;*
2. *there exists a sequence of functions $(f_n)_{n \geq 1}$, bounded in $\mathbb{D}^{1,1}$ and convergent to $f$ in $L^1 (\mathbb{P})$;*

*In such a case,*
$$|Df| (\Omega) \leq \liminf_n \mathbb{E} [|\nabla f_n|_2],$$

*and there exists a sequence $(f_n)_{n \geq 1}$ such that equality is attained.*

We remark that in the theorem above, the second condition does not mention the space $L \log^{1/2} L$: this follows by the continuous embedding discussed in Remark 2. In this sense, the extra integrability condition required in the definition of $BV$ functions is technical but natural.

## 3. A chain rule

For a composite function $f = \phi \circ g$, under certain assumptions on $\phi$ and $g$, we can conclude that $f$ is $BV$ and write an explicit formula for the integral of smooth functions with respect to any measure $D_h f$.

A chain rule for the Malliavin derivative (see [5], Proposition 1.2.3, p. 28) reads as follows: given $g_1, \ldots, g_n \in \mathbb{D}^{1,1}$ and $\phi \in C_b^1 (\mathbb{R}^n)$, then $f = \phi(g_1, \ldots, g_n) \in \mathbb{D}^{1,1}$ and
$$\nabla f = \sum_{i=1}^n \partial_i \phi(g_1, \ldots, g_n) \nabla g_i.$$

When $\phi$ is a Euclidean $BV$ function, we would like to conclude that $f$ is $BV$ (for the general theory of Euclidean $BV$ functions, see [7]). Before stating

8

some precise results, we give here a formal derivation of the chain rule in the simple case of $g_i = W(k_i)$, with $\{k_1, \ldots, k_n\}$ orthonormal in $L^2(0, T)$. Under this assumption, the joint law of $g = (W(k_1), \ldots, W(k_n))$ is the standard Gaussian law on $\mathbb{R}^n$ and we write $\rho$ for its continuous density. Given a bounded random variable $u$, there exits some Borel function $v$ such that $\mathbb{E}[u|g] = v(g)$. For brevity, we write $v(x) = \mathbb{E}[u|g = x]$. For $h \in H_0^1$, when $\phi \in C_b^1(\mathbb{R}^n)$, we integrate the chain rule above:

$$\mathbb{E}[u \langle h', \nabla f \rangle] = \sum_{i=1}^n \mathbb{E}[u \partial_i \phi(g) \langle k_1, h' \rangle_2].$$

The left member above can be replaced with $\int u D_h f$, which is defined when $f$ is $BV$. For the right member, we take the conditional expectation with respect to $g$ and we find that

$$\int u D_h f = \sum_{i=1}^n \int_{\mathbb{R}^n} \mathbb{E}[u|g = x] \langle k_1, h' \rangle_2 \rho(x) \partial_i \phi(x) \lambda^n(dx).$$

When $\phi$ is a Euclidean $BV$ function, we would like to replace $\partial_i \phi(x) \lambda^n(dx)$ with $D_i \phi(dx)$. However, we notice that the integrand in the right member above is defined $\lambda^n$ a.e. and, since $|D_i \phi|$ can be singular with respect to $\lambda^n$, it is clear that some assumptions on $u$ are necessary to give a precise meaning to the expression. In this setting, the following theorem can be established, but we omit the proof.

**Theorem 7.** *With the notation as above, if $u$ is cylindrical and continuous, then there exists a continuous representative of $\mathbb{E}[u|g = x]$ and, if $\rho.|D\phi|$ is a finite measure, then $f$ is $BV$ and*

$$\int u D_h f = \sum_{i=1}^n \int_{\mathbb{R}^n} \mathbb{E}[u|g = x] \langle k_1, h' \rangle_2 \rho(x) D_i \phi(dx),$$

*where the integrand is intended as its continuous representative.*

By a density argument, the formula above identifies the measure $Df$ as an expression of $D\phi$ and $\nabla W(k_i) = k_i$.

We turn now our attention to some chain rule formulas valid with less regularity assumptions. In what follows, we consider the case $n = 1$ only, but we do not limit ourselves to a Wiener integral: more precisely, we assume that $g$ belongs to $\mathbb{D}^{1,1}$ and moreover that its law is absolutely continuous with respect to $\lambda$, with a locally bounded density $\rho$.

Given an integrable random variable $X$, we write

$$x \mapsto \mathbb{E}[X|g = x],$$

for the equivalence class in $L^1(\rho.\lambda)$ of any Borel function $h : \mathbb{R} \to \mathbb{R}$ such that $h \circ g = P[X \mid g]$.

We introduce the following notation:

$$\xi(x) = \rho(x)\,\mathbb{E}\left[|\nabla g|_2\,|g = x\right],$$

which defines an element of $L^1(\lambda)$. We assume that $\xi$ is locally bounded and write $\xi_*(x) = \liminf_{n\to\infty} n\int_x^{x+1/n}\xi(t)\,dt$.

**Theorem 8.** *For every real $x$, the level set indicator function $f^x = I_{]x,\infty[}(g) = I_{\{g>x\}}$ is BV and $|Df^x|_2(\Omega) \le \xi_*(x)$.*

*Let $u$ be a smooth cylindrical function and $h \in H_0^1$. Then, for $\lambda$ a.e. $x$,*

$$\int u D_h f^x = \rho(x)\,\mathbb{E}\left[u\,\langle\nabla g, h'\rangle_2\,|g = x\right].$$

*Remark* 9. The left member in the identity above is a continuous function of $x$, since it coincides with $-\mathbb{E}\left[f^x\partial_h^* u\right]$ and the law of $g$ has no atoms. Therefore, the right member above admits a continuous representative.

*Proof.* For $x \in \mathbb{R}$, for $n \ge 1$, we define

$$\phi_n^x(t) = (n(t-x)\wedge 1)\vee 0,$$

and write $f_n^x = \phi_n^x \circ g$. The sequence $(f_n^x)_{n\ge 1}$ converges to $f^x$ in every $L^p(\mathbb{P})$, for $p \ge 1$. By the classical chain rule, $f_n^x$ admits a Malliavin derivative given by

$$\nabla f_n^x = n I_{]x,x+1/n[}(g)\,\nabla g.$$

Then, conditioning with respect to $g$,

$$\mathbb{E}\left[|\nabla f_n^x|_2\right] = n\int_x^{x+1/n}\xi(t)\,dt,$$

which is bounded for $n \to \infty$. By Theorem 6, the first statement is proved.

Given $u$ and $h$ as in the second statement, we write the integration-by-parts identity for every $f_n^x$:

$$-\mathbb{E}\left[\partial_h^* u f_n^x\right] = \mathbb{E}\left[u\,\langle\nabla f_n^x, h'\rangle_2\right].$$

We take the conditional expectation with respect to $g$ and integrate with respect to its law,

$$-\mathbb{E}\left[\partial_h^* u f_n^x\right] = n\int_x^{x+1/n}\rho(t)\,\mathbb{E}\left[u\,\langle\nabla g, h'\rangle_2\,|g = t\right]dt.$$

As $n \to \infty$, the right member above converges to $\rho(x)\,\mathbb{E}\left[u\,\langle\nabla g, h'\rangle_2\,|g = x\right]$, for $\lambda$ a.e. $x$, while the left member converges to $-\mathbb{E}\left[\partial_h^* u f^x\right]$, for every $x$. $\square$

Once the case of level sets is settled, we consider a $\phi \in BV_{loc}(\lambda)$ such that and $|D\phi|$ has compact support, contained in an interval $[-M, M]$, for some $M > 0$. In particular, such a $\phi$ is $\lambda$ a.e. constant for $|x|$ large (possibly with two

10

different values according to the sign of $x$). Therefore, $\phi$ differs from a Euclidean $BV$ function with compact support contained in the same interval $[-M, M]$, by adding a suitable constant function and a multiple of an indicator function of an unbounded interval. We write $C_M$ for a real number such that $\xi(x) \leq C_M$ for $\lambda$ a.e. $x \in [-M, M]$.

**Theorem 10.** *With the notation above, $f = \phi \circ g$ is $BV$, with $|Df|(\Omega) \leq C_M |D\phi|(\mathbb{R})$.*
*Let $u$ be a smooth cylindrical function and $h \in H_0^1$. Then,*

$$\int u D_h f = \int_{\mathbb{R}} \rho(x) \, \mathbb{E}[u \langle \nabla g, h' \rangle_2 | g = x] \, D\phi(dx) \tag{4}$$

*where the integrand on the right is intended as its continuous representative.*

*Proof.* We can suppose that $\phi \in BV(\lambda)$ has compact support. By the Euclidean analogue of Theorem 6 (see [7], Theorem 3.9) there exists a sequence of smooth functions with compact support $(\phi_n)_{n \geq 1}$, convergent to $\phi$ in $L^1(\lambda)$, such that the sequence of derivatives $(\phi'_n)_{n \geq 1}$ is bounded in $L^1(\lambda)$. We define $f_n = \phi_n \circ g$, so that, for some $C > 0$,

$$\mathbb{E}[|f_n - f|] = \int |\phi_n - \phi| \rho \, d\lambda \leq C \|\phi_n - \phi\|_{L^1}.$$

By the chain rule for differentiable functions, for every $n \geq 1$,

$$\nabla f_n = \phi'_n(g) \cdot \nabla g.$$

Taking expectations and conditioning with respect to $g$,

$$\mathbb{E}[|\nabla f_n|_2] = \mathbb{E}[|\phi'_n(g)| \, \mathbb{E}[|\nabla g|_2 \, | g]].$$

Integrating with respect to the law of $g$, the right member above is equal to

$$\int_{\mathbb{R}} |\phi'_n(x)| \, \xi(x) \, dx \leq C_M \int_{\mathbb{R}} |\phi'_n(x)| \, dx,$$

and we conclude that $f$ is $BV$.

To obtain the chain rule 4, we can suppose that the sequence of measure derivatives $(D\phi_n)_{n \geq 1}$ (where $D\phi_n = \phi'_n . \lambda$) converges in the duality with bounded continuous functions. It is not hard to see that equation 4 holds true with $\phi_n$ instead of $\phi$ and therefore holds true in the limit, if the integrand is intended as its continuous representative. $\qquad \square$

*Remark* 11. If $\xi$ admits a continuous representative, we obtain the stronger estimate $|Df|(\Omega) \leq \int_{\mathbb{R}} \xi(x) |D\phi|(dx)$.

In the last section we will be given a Borel function $U(x, r)$, such that, for some $s \in [0, T]$ and $\lambda$ a.e. $r \geq s$, $U(x, r)$ is a version of $\rho(x) \mathbb{E}[u \partial_r g | g = x]$, continuous in $x$. In order to obtain a chain rule in terms of $U$, we make some additional boundedness assumptions.

**Theorem 12.** *Assume that, for every real $x$, $r \mapsto U(x,r)$ belongs to $L^2(0,T)$, and the function $x \mapsto |U(x,\cdot)|_2$ is locally bounded. Then, for $\lambda$ a.e. $r \geq s$,*

$$\int u D_r f = \int_{\mathbb{R}} U(x,r) D\phi(x). \tag{5}$$

*Proof.* The following elementary fact will be useful. Let $(k_n)_{n\geq 1}$ be a bounded sequence in $L^2(0,T)$ convergent $\lambda$ a.e. to some function $k$. Then $k \in L^2(0,T)$ and the sequence converges weakly to $k$.

Because of the assumptions and the fact stated above, for every $h \in H_0^1$ with $h(r) = 0$ if $r < s$, $\int_0^T U(x,r) h'(r) dr$, is well defined and continuous as a function of $x$. Moreover, if $\psi \in C_c(\mathbb{R})$, then

$$\int_{\mathbb{R}} \psi(x) \int_0^T U(x,r) h'(r) dr dx = \int_0^T \mathbb{E}\left[u(\psi \circ g) \partial_r g\right] h'(r) dr,$$

since for every $r \geq s$, $U(x,r)$ is a version of $\rho(x) \mathbb{E}[u\partial_r g | g = x]$. We can exchange integration and expectation above, because $u(\psi \circ g)$ is bounded, and we conclude that $\int_0^T U(x,r) h'(r) dr = \rho(x) \mathbb{E}[u \langle \nabla g, h' \rangle_2 | g = x]$ for $\lambda$ a.e. $x$. This suffices to settle the case of indicator functions of level sets.

For the case of a function $\phi$, which we suppose of compact support, we consider again an approximation with smooth functions $(\phi_n)_{n\geq 1}$ with derivatives of compact support, such that $(D\phi_n)_{n\geq 1}$ weakly converges to $D\phi$. For every $n$, and every $h \in H_0^1$, with $h'(r) = 0$ if $r < s$,

$$\int_0^T h'(r) \int_{\mathbb{R}} U(x,r) D\phi_n(x) dr = \int_{\mathbb{R}} \rho(x) \mathbb{E}[u \langle \nabla g, h' \rangle_2 | g = x] D\phi_n(x),$$

Then, for $\lambda$ a.e. $r \geq s$, $\int_{\mathbb{R}} \rho(x) \mathbb{E}[u\partial_r g | g = x] D\phi_n(x)$ tends to the left member in equation (5), since the integrand is continuous. Moreover, the sequence is bounded in $L^2(s,T)$, since the integrand is locally bounded,

$$\left| \int_{\mathbb{R}} \rho(x) \mathbb{E}[u\partial_r g | g = x] D\phi_n(x) \right|_2 \leq C |D\phi_n|(\mathbb{R}).$$

As $n \to \infty$, the fact stated at the beginning of the proof applies again and we conclude that

$$\int_0^T h'(r) \left[ \int_{\mathbb{R}} U(x,r) D\phi(x) \right] dr = \int u D_h f = \int_0^T h'(r) \left[ \int u D_r f \right] dr.$$

$\square$

## 4. Ocone-Karatzas formula for $BV$ functions

*4.1. Predictable processes and projections*

Given a filtration $\mathcal{G} = (\mathcal{G}_t)_{0 \leq t \leq T}$, a $\mathcal{G}$-predictable rectangle is a subset of the product space $\Omega \times [0,T]$, of the form

$$A \times ]s,t],$$

where $A \in \mathcal{G}_s$. The $\mathcal{G}$-predictable $\sigma$-algebra of sets $\mathcal{P}_\mathcal{G}$ is, by definition, generated by the family of all $\mathcal{G}$-predictable rectangles. A stochastic process is said $\mathcal{G}$-predictable if it is measurable with respect to $\mathcal{P}_\mathcal{G}$.

Usually, one takes $\mathcal{G} = \mathcal{F}^\mathbb{P}$, the natural filtration of the Wiener process, completed with all the $\mathbb{P}$-neglegibile Borel sets, thus satisfying the so-called usual conditions, and simply speaks of predictable rectangles, predictable $\sigma$-algebra ($\mathcal{P}^\mathbb{P}$) and predictable processes. Since we are dealing with measures, in general, not absolutely continuous with respect to $\mathbb{P}$, we consider the case when $\mathcal{G} = \mathcal{F}$ is just the natural filtration of the Wiener process, and speak of *strictly* predictable rectangles, *strictly* predictable $\sigma$-algebra ($\mathcal{P}$) and *strictly* predictable processes, for the correspondent $\mathcal{F}$-predictable objects.

By a monotone class argument, every predictable process coincides $\mathbb{P} \otimes \lambda$ almost everywhere with a strictly predictable process.

A smooth process is a finite linear combination of strictly predictable processes of the form
$$f I_{]s,t]},$$
where $f$ is smooth. The strict predictability implies that $f$ can be written as $\phi(\Delta_J W)$ for some $\Delta_J W$, as introduced in (1), such that $s_n \leq s$. By a direct approximation of cylindrical functions with smooth functions and a monotone class argument, smooth processes are dense in every $L^p(\Omega \times [0,T], \mathcal{P}, \mu)$, for $1 \leq p < \infty$, where $\mu$ is any positive finite measure on $\mathcal{P}$. Therefore, if we are given two real measures $\mu$, $\nu$ on $\mathcal{P}$ such that for every smooth process $F$, $\int F d\mu = \int F d\nu$, then $\mu = \nu$.

The classical Ocone-Karatzas formula contains the stochastic integral of a predictable process $H = (H_s)_{0 \leq s \leq T}$, which can be correctly defined as the density of the measure $\nabla f. (\mathbb{P} \otimes \lambda)$ restricted to the strictly predictable $\sigma$-algebra $\mathcal{P}$, with respect to $\mathbb{P} \otimes \lambda$. The process $H$ is also identified by the condition $H_s = \mathbb{E}[\partial_s f | \mathcal{F}_s]$, for $\lambda$ a.e. $s \in [0,T]$. We can see $H_s$ as the density of $\partial_s f. \mathbb{P}$, restricted to $\mathcal{F}_s$, with respect to $\mathbb{P}$ (also restricted).

Given a $BV$ function $f$, Theorem 16 will prove that its measure derivative $Df$, as a measure on the product space, when restricted to $\mathcal{P}$, becomes absolutely continuous with respect to $P \otimes \lambda$, with a strictly predictable density $H = (H_s)_{0 \leq s \leq T}$. The next lemma shows that we can still see $H_s$ as the density of the measure $D_s f$, restricted to $\mathcal{F}_s$, with respect to $\mathbb{P}$. We state it for a general real measure $\mu$ on the product space $(\Omega \times [0,T], \mathcal{A} \otimes \mathcal{B}(0,T))$ of the form $\mu = \sigma. (\nu \otimes \lambda)$, where $\nu$ is a positive finite measure on $(\Omega, \mathcal{A})$.

**Proposition 13.** *With the notation above, if $\mu_{|\mathcal{P}}$ is absolutely continuous with respect to $(P \otimes \lambda)_{|\mathcal{P}}$, with a strictly predictable density $H = (H_s)_{0 \leq s \leq T}$, then, for $\lambda$ a.e. $s \in [0,T]$,*
$$(\sigma(s).\nu)_{|\mathcal{F}_s} = H_s. P_{|\mathcal{F}_s}.$$

*Proof.* For every $n \geq 1$, there exists a countable family of bounded continous function $(f_k^n)_{k \geq 1}$, defined on $\mathbb{R}^n$, such that every $f \in C_c(\mathbb{R}^n)$ is the pointwise limit of an appropriate sequence with elements from that family. Therefore, for

$\lambda$ a.e. $s \in [0, T]$, it will suffice to prove that, for all $n \geq 1$, for every $\Delta_J W$, with $J = (t_0, \ldots, t_n)$ with $t_n \leq s$ and every $k \geq 1$,

$$\int f_k^n (\Delta_J W) \sigma(s) \, d\nu = \mathbb{E}\left[ f_k^n (\Delta_J W) H_s \right].$$

Moreover, by continuity, we can consider only the case with each $t_i$ rational. Therefore, the thesis is equivalent to prove that for every $n$, every $\Delta_J W$, with each $t_i$ rational, and every $k \geq 1$, for $\lambda$ a.e. $s \geq t_n$, the identity above holds true. The hypothesis implies that for every $t \geq t_n$,

$$\int_{t_n}^t \int f_k^n (\Delta_J W) \sigma(s) \, d\nu ds = \int_{t_n}^t \mathbb{E}\left[ f_k^n (\Delta_J W) H_s \right] ds.$$

since it is the integral of a stricly predictable process. Therefore the two integrands must coincide for $\lambda$ a.e $s \geq t_n$. $\qquad\square$

### 4.2. Statement and proof of the main result

For the proof of the extended Ocone-Karatzas formula we will make use of the following result, which contains Itô's representation theorem and a version of Ocone-Karatzas formula for smooth functions, which can be shown directly by an application of Itô's formula.

**Lemma 14.** *For $f \in L^1(\mathbb{P})$, there exists a unique process $F \in \mathcal{M}_{loc}^2$ such that*

$$f = \mathbb{E}[f] + \int_0^T F_s dW_s. \tag{6}$$

*If $f$ is smooth, then $F_s = \mathbb{E}[\partial_s f | \mathcal{F}_s]$, for $\lambda$ a.e. $s \in [0, T]$.*

*Remark* 15. The space $\mathcal{M}_{loc}^2$ consists of all the equivalence classes of predictable processes $F$, such that

$$\int_0^T F_s^2 ds < \infty, \quad \mathbb{P} \text{ almost surely.}$$

For such an $F$, there exists a sequence of predictable stopping times $(\tau_n)_{n \geq 1}$, increasing towards $T$, such that, for every $n$, the predictable process $F_n = F I_{[0, \tau_n[}$ is square integrable (and so it belongs to $\mathcal{M}^2$).

**Theorem 16.** *Let $f$ be a BV function. The measure $Df$, restricted to the strictly predictable $\sigma$-algebra $\mathcal{P}$, is absolutely continuous with respect to $\mathbb{P} \otimes \lambda$. If $H$ is a version of the density, then*

$$f = \mathbb{E}[f] + \int_0^T H_s dW_s.$$

*Remark* 17. By Proposition 13, we can write, for $\lambda$ a.e. $s \in [0, T]$:

$$H_s = \frac{d(D_s f)_{|\mathcal{F}_s}}{dP_{|\mathcal{F}_s}}.$$

14

We provide here a shorter proof for case $f \in BV \cap L^2(\mathbb{P})$.

*Proof.* There is no loss in generality if we suppose that $\mathbb{E}[f] = 0$. Since $f$ is square integrable, the process $F$ in (6) belongs to $\mathcal{M}^2$ and therefore $F.\mathbb{P} \otimes \lambda$ defines a real measure on the product space. It is sufficient to show that the integrals of smooth processes, taken with respect to $Df$ and $F.\mathbb{P} \otimes \lambda$, coincide.

Given $0 \le s < t \le T$ and a smooth function $g$, measurable with respect to $\mathcal{F}_s$, the Wiener-Itô isometry implies that

$$\mathbb{E}\left[f \int_s^t g dW_r\right] = \mathbb{E}\left[\int_0^T F_r dW_r \int_s^t g dW_r\right] = \mathbb{E}\left[\int_s^t g F_r dr\right].$$

Since $g$ is $\mathcal{F}_s$-measurable, $\int_s^t g dW_r = -\partial_h^* g$, where $h \in H_0^1$ is such that $h' = I_{]s,t]}$. By the definition of $Df$, the left member above is equal to

$$\int g I_{]s,t]} \, dDf.$$

By linearity, we conclude. $\square$

The next lemma shows that we gain integrability for $F$ in (6), when $f$ is $BV$.

**Lemma 18.** *If $f$ is $BV$ and $F$ is the process in (6), then*

$$\int_0^T \mathbb{E}[|F_s|] \, ds \le T^{1/2} |Df|(\Omega).$$

*Remark* 19. If $f \in L \log L$, then, by the continuous embedding of this Orlicz space into $H^1$ (see e.g. [8], Exercise 1.16, p. 58) and by the equivalence of the $H^1$ norm with $\mathbb{E}[|f| + |F|_2]$, we conclude that $F$ is integrable. However, it is not clear if such a conclusion holds true for $f \in L \log^{1/2} L$: we proceed therefore generalizing some ideas from [4].

*Proof.* If $f$ is smooth, then $F = \mathbb{E}[\partial_s f | \mathcal{F}_s]$, so that $|F_s| \le \mathbb{E}[|\partial_s f| \,|\mathcal{F}_s]$. By Fubini's theorem and Cauchy's inequality,

$$\mathbb{E}\left[\int_0^T |F_s| \, ds\right] \le \mathbb{E}\left[\int_0^T |\partial_s f| \, ds\right] \le T^{1/2} \mathbb{E}[|\nabla f|_2].$$

Given a $BV$ function $f$, we consider a sequence of smooth functions $(f_n)_{n \ge 1}$, convergent to $f$ in $L^1(\mathbb{P})$, such that $(\mathbb{E}[|\nabla f_n|_2])_{n \ge 1}$ converges to $|Df|(\Omega)$. Such a sequence exists by an application of Theorem 6 and a diagonal argument.

For every $n \ge 1$, we write $F_n$ for the process that represents $f_n$, and $M^n$ for the $\mathcal{F}^{\mathbb{P}}$-martingale

$$M_t^n = \mathbb{E}\left[f_n | \mathcal{F}_t^{\mathbb{P}}\right] = \mathbb{E}[f_n] + \int_0^t F_s^n dW_s.$$

15

We also write $M_t = \mathbb{E}\left[f|\mathcal{F}_t^{\mathbb{P}}\right]$, and $\Delta M_n^* = \sup_{0 \leq t \leq T} |M_t^n - M_t|$. By Doob's maximal inequality, we have

$$\mathbb{P}\left(\Delta M_n^* > \epsilon\right) \leq \frac{1}{\epsilon}\mathbb{E}\left[|M_T^n - M_T|\right],$$

for every $\epsilon > 0$. By the Burkholder-Gundy good-$\lambda$-inequality (see [9], and [4]),

$$\mathbb{P}\left(\int_0^T |F_s^n - F_s|^2\,ds > 4\lambda^2, \Delta M_n^* \leq \delta\lambda\right) \leq \delta^2\mathbb{P}\left(\int_0^T |F_s^n - F_s|^2\,ds > \lambda^2\right),$$

for every $\lambda > 0$ and $\delta \in (0,1)$. This leads immediately to the inequality

$$\mathbb{P}\left(\int_0^T |F_s^n - F_s|^2\,ds > 4\lambda^2\right) \leq \delta^2 + \frac{1}{\delta\lambda}\mathbb{E}\left[|M_T^n - M_T|\right],$$

which implies that

$$\int_0^T |F_s^n - F_s|^2\,ds \to 0$$

in probability, and without loss of generality, $\mathbb{P}$ almost surely. By Cauchy's inequality, $\int |F_s^n|\,ds$ converges to $\int |F_s|\,ds$, $\mathbb{P}$ almost surely so, by Fatou's lemma, we conclude that

$$\mathbb{E}\left[\int_0^T |F_s|\,ds\right] \leq T^{1/2}\liminf_n \mathbb{E}\left[|\nabla f_n|_2\right] = T^{1/2}\,|Df|\,(\Omega).$$

$\square$

*Proof.* (General case.) We can argue as in the $L^2(\mathbb{P})$ case, since by the lemma above, $F.(\mathbb{P} \otimes \lambda)$ defines a real measure on the product space. It is therefore sufficient to show that the integrals of smooth processes, taken with respect to $Df$ and $F.\mathbb{P} \otimes \lambda$, coincide. We take $0 < s < t \leq T$ and a smooth $g$, measurable with respect to $\mathcal{F}_s$.

Let $(\tau_n)_{n \geq 1}$ be a sequence of predictable stopping times increasing to $T$, such that $F^n = I_{[0,\tau_n[}F \in \mathcal{M}^2$. Then,

$$f_n = \mathbb{E}\left[f|\mathcal{F}_{\tau_n}^{\mathbb{P}}\right] = \int_0^T F_r^n\,dW_r,$$

so that $(f_n)_{n \geq 1}$ an $\left(\mathcal{F}_{\tau_n}^{\mathbb{P}}\right)_{n \geq 1}$-martingale, closed by $f$, in $L\log^{1/2}L$. By Lemma 3, $(f_n)_{n \geq 1}$ converges to $f$ in $L\log^{1/2}L$.

For every $n \geq 1$, by the Wiener-Itô isometry,

$$\mathbb{E}\left[f_n\,g\left(W_t - W_s\right)\right] = \mathbb{E}\left[\int_0^T g I_{]s,t]}(r)\,F_r^n\,dr\right]. \tag{7}$$

16

By inequality (2), as $n \to \infty$, the left member converges to

$$\mathbb{E}\left[fg\left(W_t - W_s\right)\right] = \int g I_{]s,t]} \, dDf.$$

For the member on the right, we use Lebesgue's dominated convergence theorem, since $gF \in L^1\left(\Omega \times [0,T], \mathbb{P} \otimes \lambda\right)$. Therefore, identity (7) holds true with $f$ in place of $f_n$, and we conclude as we did for the $L^2\left(\mathbb{P}\right)$ case. $\square$

## 5. Applications

We conclude by showing how the extended Ocone-Karatzas formula, together with the chain rule, allows us to find explicit representations. We consider first the case of $f = \phi\left(W\left(k\right)\right)$ and then the case of $f = \phi\left(M\right)$ where $M = \sup_{s \leq T} W_s$, with $\phi$ as introduced in Section 3.

### 5.1. The case of cylindrical random variables

Given $k \in L^2\left(0, T\right)$, with $|k|_2 = 1$, we choose a representative such that $k\left(s\right) = 0$ if $\left|I_{]s,T]}k\right|_2 = 0$, for every $s \in [0, T]$, and write

$$K\left(x, s\right) = k\left(s\right) \frac{\exp\left\{-x^2 / \left(2\left|I_{]s,T]}k\right|_2^2\right)\right\}}{\sqrt{2\pi}\left|I_{]s,T]}k\right|_2},$$

where we define $K\left(x, s\right) = 0$ if $k\left(s\right) = 0$.

**Proposition 20.** *Let* $f = \phi\left(W\left(k\right)\right)$, *with the notation as above. Then*

$$f = \mathbb{E}\left[f\right] + \int_0^T \left[\int_{\mathbb{R}} K\left(x - W\left(I_{]0,s]}k\right), s\right) D\phi\left(dx\right)\right] dW_s.$$

*Proof.* It is sufficient to prove that, for all $s \in [0, T]$ and every smooth function $u$, measurable with respect to $\mathcal{F}_s$, then

$$\frac{\exp\left\{-x^2/2\right\}}{\sqrt{2\pi}} \mathbb{E}\left[uk\left(s\right) \mid W\left(k\right) = x\right] = \mathbb{E}\left[u\, K\left(x - W\left(I_{]0,s]}k\right), s\right)\right].$$

Indeed, we define $U\left(r, x\right)$ as the right member above if $r \geq s$ and $U\left(r, x\right) = 0$ for $r < s$. Since the right member above is a continuous function of $x$ and the inequality

$$\mathbb{E}\left[uk\left(s\right) \mid W\left(k\right) = x\right] \leq \|u\|\left|k\left(s\right)\right|$$

holds true for all $s \in [0, T]$, we apply Theorem 12, Fubini's theorem and we find that

$$\mathbb{E}\left[u \int_{\mathbb{R}} K\left(x - W\left(I_{]0,s]}k\right), s\right) D\phi\left(dx\right)\right] = \int u D_s f.$$

for $\lambda$ a.e. $s \in [0, T]$ and every smooth function $u$ as above. By Remark 17, we conclude.

Therefore, given a continuous function $\psi$, with compact support, we have to show that

$$\mathbb{E}\left[\psi\left(W\left(k\right)\right)uk\left(s\right)\right]=\int_{\mathbb{R}}\psi\left(x\right)\mathbb{E}\left[\,u\,K\left(x-W\left(I_{]0,s]}k\right),s\right)\,\right]dx.$$

We take the conditional expectation with respect to $\mathcal{F}_s$,

$$\mathbb{E}\left[\psi\left(W\left(k\right)\right)uk\left(s\right)\right]=\mathbb{E}\left[u\int_{\mathbb{R}}\psi\left(W\left(kI_{]0,s]}\right)+y\right)K\left(y,s\right)dy\right].$$

We conclude thanks to the change of variables $x=W\left(kI_{]0,s]}\right)+y$ and exchanging integration and expectation. $\qquad\square$

*Example* 1. We consider $f=I_{\{W_t\geq0\}}$, for some $0<t\leq T$. Since $DI_{\left\{\sqrt{t}x\geq0\right\}}=\sqrt{t}\delta_0$, we find the representation

$$I_{\{W_t\geq0\}}=\frac{1}{2}+\int_0^t\frac{\exp\left\{-W_s^2/2\left(t-s\right)\right\}}{\sqrt{2\pi\left(t-s\right)}}dW_s.\tag{8}$$

*Remark* 21. Formula (8) above is well known (e.g. [10], p. 68) and there exists even an alternative approach (see [11]), which produces similar formulas, valid for bounded cylindrical random variables.

Indeed, the following observation shows that the two approaches are, in this case, equivalent. For every $s$, $K\left(\cdot,s\right)\in C_b^1\left(\mathbb{R}\right)$, so we can integrate-by-parts on $\mathbb{R}$ and obtain the representation

$$f=\mathbb{E}\left[f\right]+\int_0^T\left[\int_{\mathbb{R}}-\phi\left(x\right)\partial_xK\left(x-W\left(I_{]0,s]}k\right),s\right)dx\right]dW_s,$$

which is a special case of Theorem 3.4 in [11].

We remark that, in the next application, such an integration-by-parts will fail in general, since the correspondent $K$ is less regular.

*5.2. The case of the maximum of the Wiener process*

Given $0\leq s<t\leq T$, we write $M_{[s,t]}=\sup_{s\leq r\leq t}\left(W_r-W_s\right)$, so that,

$$M=M_{[0,T]}=M_{[0,s]}\vee\left(W_s+M_{[s,T]}\right)\tag{9}$$

and $M_{[s,T]}$ is independent of $\mathcal{F}_s$, with a.c. law (with respect to $\lambda$), of density

$$m_{T-s}\left(x\right)=2\frac{\exp\left\{-x^2/2\left(T-s\right)\right\}}{\sqrt{2\pi\left(T-s\right)}}I_{\{x>0\}}.$$

**Proposition 22.** *Let* $f=\phi\left(M\right)$, *with the notation as above. Then*

$$f=\mathbb{E}\left[f\right]+\int_0^T\left[\int_{\mathbb{R}}m_{T-s}\left(x-W_s\right)I_{\left\{x>M_{[0,s]}\right\}}D\phi\left(dx\right)\right]dW_s.\tag{10}$$

18

*Remark* 23. It is known (see e.g. [5], Proposition 2.1.10, p. 109) that $M$ admits a Malliavin derivative, $\nabla M = I_{[0,\sigma[}$, where

$$\sigma = \inf\left\{t : W_t \geq W_s, \ \forall 0 \leq s \leq T\right\}.$$

We note that $\partial_s M$ is the indicator function of $\left\{M_{[0,s]} < W_s + M_{[s,T]}\right\}$.

*Proof.* It is sufficient to prove that, for all $s \in ]0, T[$ and every smooth function $u$, measurable with respect to $\mathcal{F}_s$, then

$$m_T\left(x\right)\mathbb{E}\left[u\partial_s M \,|M = x\right] = \mathbb{E}\left[u\,m_{T-s}\left(x - W_s\right)I_{\left\{x > M_{[0,s]}\right\}}\right].$$

Indeed, we define $U\left(r, x\right)$ as the right member above if $r \geq s$ and $U\left(r, x\right) = 0$ for $r < s$. Since the right member above is a continuous function of $x$ and the inequality

$$\mathbb{E}\left[u\partial_s M \ |M = x\right] \leq \|u\|$$

holds true for all $s \in ]0, T[$, we conclude as we did for cylindrical functions.

Therefore, given a continuous function $\psi$, with compact support, we have to show that

$$\mathbb{E}\left[\psi\left(M\right)u\partial_s M\right] = \int_{\mathbb{R}} \psi\left(x\right)\mathbb{E}\left[u\,m_{T-s}\left(x - W_s\right)I_{\left\{x > M_{[0,s]}\right\}}\right]dx.$$

We take the conditional expectation with respect to $\mathcal{F}_s$ and use the expression for $M$ in (9) $\partial_s M$ in Remark 23,

$$\mathbb{E}\left[\psi\left(M\right)u\partial_s M\right] = \mathbb{E}\left[u\int_{\mathbb{R}} \psi\left(W_s + y\right)I_{\left\{W_s + y > M_{[0,s]}\right\}}m_{T-s}\left(y\right)dy\right].$$

With a change of variables $x = W_s + y$ and exchanging integration and expectation, we conclude. $\square$

*Example* 2. We consider the function $I_{\{M \geq y\}}$, for some $y > 0$. Since $DI_{\{x \geq y\}} = \delta_y$, we find the representation

$$I_{\{M \geq y\}} = \mathbb{P}\left(M \geq y\right) + \int_0^{\tau_y} m_{T-s}\left(y - W_s\right)dW_s,$$

where $\tau_y = \inf\left\{0 \leq s \leq T : W_s = y\right\}$ is the time of the first visit at $y$.

## Acknowledgements

# References

[1] M. Fukushima, *BV* functions and distorted Ornstein-Uhlenbeck processes over the abstract Wiener space, Journal of Functional Analysis 174 (2000) 227–249.

[2] M. F. M. Hino, On the space of *BV* functions and a related stochastic calculus in infinite dimensions, Journal of Functional Analysis 183 (2001) 245–268.

[3] L. Ambrosio, M. M. Jr., S. Maniglia, D. Pallara, *BV* functions in abstract Wiener spaces, Journal of Functional Analysis 258 (2010) 785–813.

[4] I. Karatzas, D. L. Ocone, J. Li, An extension of Clark's formula, Stochastics and Stochastics Reports 37 (1991) 127–131.

[5] D. Nualart, The Malliavin calculus and related topics, Probability and its applications, Springer-Verlag, 1995.

[6] M. Ledoux, Isoperimetry and gaussian analysis, in: Lectures on Probability Theory and Statistics, Vol. 1648 of Lecture Notes in Mathematics, Springer Berlin / Heidelberg, 1996, pp. 165–294, 10.1007/BFb0095676.

[7] L. Ambrosio, N. Fusco, D. Pallara, Functions of bounded variation and free discontinuity problems, Oxford mathematical monographs, Clarendon Press, 2000.

[8] D. Revuz, M. Yor, Continuous martingales and Brownian motion, Grundlehren der mathematischen Wissenschaften, Springer, 1999.

[9] D. L. Burkholder, Distribution function inequalities for martingales, The Annals of Probability 1 (1) (1973) pp. 19–42.

[10] A. Üstünel, An introduction to analysis on Wiener space, Lecture notes in mathematics, Springer, 1995.

[11] K. Aase, B. ksendal, J. Ube, Using the donsker delta function to compute hedging strategies, Potential Analysis 14 (2001) 351–374.